Psychological Monographs
General and Applied

No. 607
1965

A Quantitative Extension of Heider's Theory of Cognitive Balance Applied to Interpersonal Perception and Self-Esteem

By

William M. Wiest

Reed College

Price $1.00

Vol. 79
No. 14

William M. Wiest

Edited by Gregory A. Kimble
Published by the American Psychological Association, Inc.

Psychological Monographs:
General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

Psychological Monographs does not maintain a fixed board of consulting editors. A large group of psychologists has agreed to help in the evaluation of manuscripts. The names of those who serve in this capacity will be published in the last number of each volume.

Manuscripts and correspondence on editorial matters should be sent to the Editor: Gregory A. Kimble, Department of Psychology, Duke University, Durham, North Carolina 27706. *Psychological Monographs* publishes comprehensive experimental investigations and other psychological studies which do not lend themselves to adequate presentation as journal articles. A *Monograph* should be completely understandable in itself. Introductory materials, including historical and theoretical background, should be developed in enough detail to preclude the necessity for consulting other references to understand the *Monograph*. Thorough description of procedures and results and discussions of implications are encouraged. On the other hand, the content of the *Monograph* should include only what is relevant. Tables, graphs, and appendixes which present material not essential to adequate understanding of the study may be made available through the American Documentation Institute. The procedures for making use of this facility are described in the APA *Publication Manual*. Manuscripts for publication as *Monographs* should adhere to the standards given in the APA *Publication Manual*. A statement of policies which are specific to *Psychological Monographs* appeared in the *American Psychologist*, 1964, **19**, 284–285. Publication in *Psychological Monographs* is free of cost to the author except in cases where author's alterations are made in page proofs.

The following policies govern reprinting of materials copyrighted in APA journals: (*a*) to require approval to reprint for tables and figures, and for text only if more than 500 words [total from one article] in length; (*b*) to continue approval to reprint, contingent on the author's approval, for articles reprinted in whole or in major part; (*c*) to negotiate where possible the dedication of royalties from commercial publishers to the American Psychological Foundation.

Manuscripts must be accompanied by an abstract of 100–200 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*, Detailed instructions for the preparation of abstracts appeared in the *American Psychologist* (1961, **16**, 833), or they may be obtained from the Editor or from the APA Central Office.

Helen Orr	Cornelia McConnell
Managing Editor	Editorial Assistant

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1200 Seventeenth Street, N. W., Washington, D. C. 20036. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

Vol. 79, No. 14 Whole No. 607, 1965

Psychological Monographs: General and Applied

A QUANTITATIVE EXTENSION OF HEIDER'S THEORY OF COGNITIVE BALANCE APPLIED TO INTERPERSONAL PERCEPTION AND SELF-ESTEEM[1]

WILLIAM M. WIEST

Reed College

Theories of cognitive balance (with the partial exception of Osgood & Tannenbaum's congruity model) assume only dichotomous measurement of relations among elements. To eliminate this restriction, a quantitative extension of Heider's theory is presented that enables systematic treatment of cognitive structures in which all relations among elements are continua rather than dichotomies. Sociometric-like ratings of classmates by 415 children (Grades 5, 6, and 7) support the quantitative model's prediction that (a) the magnitude and sign of correlation between S's feeling toward a set of acquaintances, and his perception of a focal person's evaluation of these same persons, is a linear function of S's attraction to the focal other ($p < .0001$), and (b) the extent to which S believes his evaluations of others are reciprocated by them is a positive function of his self-esteem ($p < .0005$).

RECENT theoretical formulations by theorists such as Abelson and Rosenberg (1958), Cartwright and Harary (1956), Festinger (1957), Heider (1946, 1958), McGuire (1960a, 1960b), Newcomb (1953, 1961), and Osgood and Tannenbaum (1955) focus attention upon the relative congruity or consistency among a person's cognitions. While these formulations differ from one another both in rigor and in the types of situations to which they typically are applied, they agree in emphasizing the interdependence among elements of a cognitive structure. Each theory defines certain sets of relations among cognitive elements as balanced, congruous, consistent, or consonant, and each postulates that states of imbalance tend to become resolved into balanced states. Data from a wide variety of settings attest to the predictive power of these theories of cognitive consistency (Brehm & Cohen, 1962; Burdick & Burnes, 1958; Cohen, 1960; Horowitz, Lyons, & Perlmutter, 1951; Jordan, 1953; Kogan & Tagiuri, 1958; Morrissette, 1958; Newcomb, 1953, 1961; Osgood, 1960; Rosenberg, Hovland, McGuire, Abelson, & Brehm, 1960; Runkel, 1956; Sampson & Insko, 1964).

One limitation shared by most theories of cognitive balance is that they are restricted to the special case in which the relations among cognitive elements are measured dichotomously. While Abelson and Rosenberg's (1958) model and Cartwright and Harary's (1956) generalization of Heider's theory make it possible to define the degree of balance of structures with more than three elements, these contributions leave unsolved the problem of how to deal with relationships of different strength among elements. Osgood and Tannenbaum (1955), treating some relations among elements quantitatively, have made a significant initial step toward adequate handling of this issue. Some features of their model of attitudinal congruity can be treated as special cases of the quantitative extension of Heider's theory proposed here.

[1] This paper is adapted from a dissertation submitted to the University of California, Berkeley, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Abbreviated reports of Study I and of Study II were presented at the 1963 annual meetings of the American Psychological Association and the Western Psychological Association, respectively. The author is especially grateful to M. Brewster Smith, chairman of the doctoral committee, for his support and advice in every phase of the study; John A. Clausen and Edward E. Sampson also offered valuable assistance. The personnel and children of the Berkeley school system are thanked for their role in making the study possible.

Because current theories of cognitive balance lack a clear rationale for defining the degree of balance in structures whose elements have variable amounts of similarity or attraction (or other relation) to one another, they can not reflect accurately the degree of balance found in empirically measured cognitions. Further, as is true of all nonquantitative formulations, these theories place severe restriction on the extent to which interaction between theory and data can aid in the development of more powerful scales of measurement.

To obviate these limitations of nonquantitative theory, a more general model of cognitive balance is here proposed. The model is designed to explicate a balance theory for the general case in which the relations among elements in the cognitive structures are continua rather than dichotomies. The present formulation focuses primarily upon structures having no more than three relations among the elements. However, preliminary theoretical and empirical attempts by the author to generalize the model to a larger number of relations (dimensions) appear promising.

The model is referred to as an extension of Heider's theory because it derives primarily from Heider's work and adopts his terminology and because of Heider's apparent precedence in the explicit statement of balance theory. Nevertheless, the model is closely related to other theories of cognitive consistency.

Several features of Heider's (1958) theory relevant to the present formulation may be noted. The theory typically deals with cognitive structures having two or three elements—p, a person or perceiver; o, some other person; and x, an impersonal object or thing. While various kinds of relations among elements can be defined, Heider distinguishes two classes—sentiment, or liking relations, L, and similarity, or unit relations, U. Each of the relations may have either a positive or a negative sign. For example, p can either like or dislike o, p can think that o is or is not responsible for x, etc.

While Heider treats L and U relations as functionally equivalent, Cartwright and Harary (1956) have noted that ~L means "dislikes," the opposite of "likes," while ~U generally is understood to mean "not related or similar to," the complement of "related to." The negation of the U relation implies the absence of a relation and therefore might better be represented by zero than by a negative sign. The fact that negative U relations result in "vacuously" balanced structures (Cartwright & Harary, 1956) is explicable in terms of the theory proposed here.

The remaining discussion focuses primarily on the more general affective relation, L. Thus, pLo represents the variable, p's degree of attraction to o, pLx is p's evaluation of x, and oLx is p's perception of o's feeling toward x. With relations between pairs of elements measured dichotomously, the eight p-o-x configurations shown in Figure 1 can be generated. Those structures with all positive or two negative signs are defined as balanced, and those with one or three negative signs as imbalanced. While Heider (1958, pp. 203, 206) feels there is some ambiguity in the case of three negative relations, the rationale provided by graph theory (Cartwright & Harary, 1956) justifies treating this case as imbalanced.

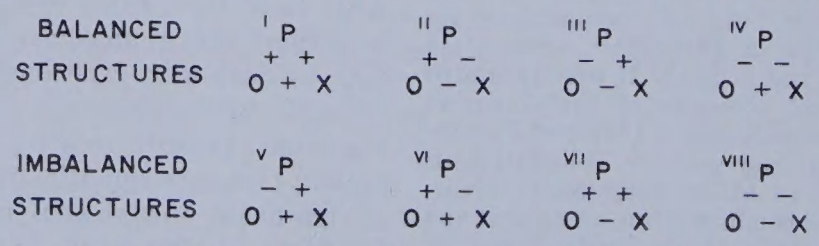

FIG. 1. Balanced and imbalanced structures for the case of three elements related positively or negatively. (The signs between p and o and between p and x represent p's evaluation of (or more generally, p's relation to) o and x, respectively, and the sign between o and x represents p's perception of the relation between o and x. The three specific liking relations are referred to as pLo, pLx, and oLx.)

Because the theory implies that imbalanced states generate forces toward restoration of balance, fewer imbalanced than balanced structures should be observed empirically. This hypothesis is strongly supported by Kogan and Tagiuri's (1958) study of the perception of interpersonal preferences among members of small groups, a study in which another person, q, substituted for x in the p-o-x triad. Studying two balanced structures (Cases I and IV in Figure 1) and two imbalanced ones (Cases V and VI), these investigators found that balanced cognitive structures occur far in excess of chance and also in excess of the degree of balance of the actual preference network. Imbalanced structures, on the other hand, occur significantly less often than would be expected by chance.

The Quantitative Balance Model

In the quantitative version of Heider's theory each relation, pLo, pLx, and oLx, is considered a bipolar dimension conceptually independent of the other two. Each dimension can be identified with a quantitative scale measuring the degree of (felt or perceived) attraction or antipathy between pairs of elements. These three relations, taken as variables, are viewed as defining a three-dimensional space in which each possible point represents a configuration of values on the three relations (see Figure 2).

Heider's four balanced structures can be seen on the left side of Figure 2, located at the four corners of the cube, I, II, III, and IV. For example, the corner point labeled "I," represents that configuration in which p is highly attracted to both o and x and in which o is also perceived to have a high degree of liking for x.

All points in the space that fall along a straight line joining any of Heider's four balanced structures, are defined as balanced. The six possible straight lines that can be defined in this manner yield a tetrahedron that lies within the cube. The exact form and position of the tetrahedron is easily seen by examining the right hand side of Figure 2, where the tetrahedron is portrayed by itself. It is assumed that all balanced configurations (points) are on the surface of or within the tetrahedron. However, the closer a point is to the center of the tetrahedron, the more does the characterization, vacuously balanced, apply; this is because approaching the center of the tetrahedron is equivalent to approaching a zero value on all the variables (i.e., it is equivalent to approaching the *absence* of any attraction, repulsion, or similarity relation among the elements in the structure). Points outside the tetrahedron are imbalanced, and they are more imbalanced the greater their distance from the nearest side of the tetrahedron. Thus, the four corners of the cube maximally distant from the nearest surface of the tetrahedron are coordinated to the four imbalanced structures of Figure 1.

Of course, it is of interest to specify those points in the space which are balanced (or balanced to varying degrees) because of the hypothesis that balanced states are relatively stable equilibria toward which imbalanced configurations tend to move. Thus, to postulate that a particular figure defines the limits within which all balanced structures lie is to postulate that empirically measured configurations will tend to be contained within these boundaries. In short, balance theory in this context predicts certain features of the distribution of points in the space.

One method for translating these implications of the model into empirical operations yields a set of predictions about

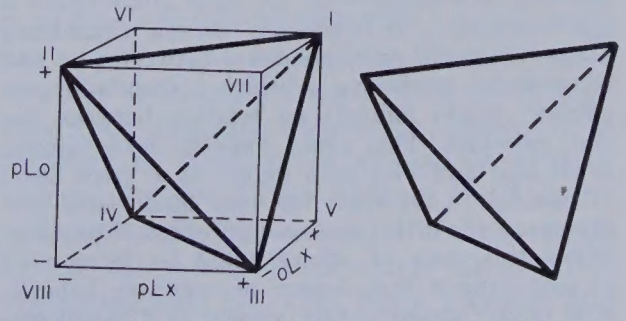

Fig. 2. Three-dimensional space defined by the relations, pLo, pLx, and oLx. (The heavy lines within the cube represent the six edges of a tetrahedron lying inside the cube; examining the tetrahedron alone to the right of the cube may facilitate seeing the tetrahedron within the cube. All balanced configurations [points] are on the surface of or within the tetrahedron.)

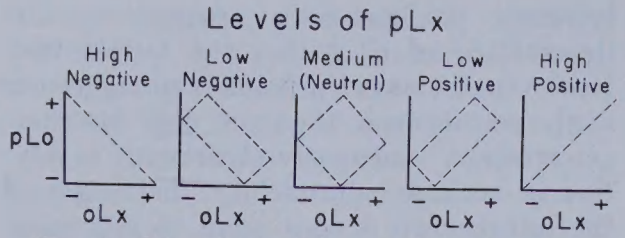

FIG. 3. Hypothetical boundaries of balanced points at five different levels of pLx, obtained by "slicing" the cube of Figure 2 at five places along the dimension, pLx.

the magnitude of covariation (or association) between any two variables as a function of the third variable. Such an interpretation is illustrated in Figure 3, which shows the cube of Figure 2 "sliced" in five places along the dimension, pLx. The same "picture" would emerge if the cube were cut in either of the other two dimensions.

If empirically measured configurations tend to lie within the boundaries hypothesized for the extended balance theory, then the direction and degree of correlation between pLo and oLx should be directly related to the value of pLx. It is apparent that the model in this form can be tested by calculating the degree of correlation[2] between any two variables with the third held constant at a number of different levels. Thus, the predictions of the model can be stated formally as a set of monotonic functional relations, of which only Case 3 is shown graphically in Figure 3;

$$1.\ r_{\mathrm{pLo,\ pLx}} = f(\mathrm{oLx}),$$

$$2.\ r_{\mathrm{pLx,\ oLx}} = f(\mathrm{pLo}),$$

$$3.\ r_{\mathrm{pLo,\ oLx}} = f(\mathrm{pLx}).$$

In verbal translation the notation implies that the extent to which the intensity of any two relations covaries is a function of the third relation. For example (Case 3), the more intensely p feels toward x, the greater will be the covariation between his feelings toward o and his perception of o's feelings toward x.

The predictions can be stated more rigorously, if the measuring instruments warrant, as a set of corresponding linear equations;

$$1.\ r_{\mathrm{pLo,\ pLx}} = k + w(\mathrm{oLx}),$$

$$2.\ r_{\mathrm{pLx,\ oLx}} = k + w(\mathrm{pLo}),$$

$$3.\ r_{\mathrm{pLo,\ oLx}} = k + w(\mathrm{pLx}),$$

where k = some constant ($k = 0$ if the implications of the model are taken literally) and w = a weight whose value depends upon the scale units of the variable being weighted (e.g., $w = 1$ if $+1$ represents the maximum amount of liking and -1 represents maximum disliking). Stated verbally and with some loss in precision, several implications of Equation 3, for example, are: (*a*) The more p likes x, the greater will be the degree of positive correlation between his liking for o and his perception of o's liking for x. (*b*) The more p dislikes x, the greater will be the degree of negative correlation between his liking for o and his perception of o's liking for x. (*c*) If p feels indifferent toward (neither likes nor dislikes) x, there will be no relation between his liking for x and his perception of o's liking for x.

The more rigorous version of the model (as a set of linear equations) can be tested adequately only if rather strong assumptions are made about the properties of the

[2] An average difference score, $\bar{D}$, or some variety of difference score (cf. Osgood, Suci, & Tannenbaum, 1957), providing a measure of the correspondence between, for example, the pLo values and the oLx values at different levels of pLx, might appear to be as good a dependent variable as the correlation coefficient for purposes of testing the model. However, the $\bar{D}$ index is less adequate than a correlation coefficient in a number of respects. First, unlike a correlation coefficient, the numerical value of $\bar{D}$ is not independent of the unit of measurement. Of more importance is the fact that a $\bar{D}$ index may be low (indicating a high degree of correspondence between pLo and oLx) under conditions in which a correlation coefficient would indicate no relation between the two variables. This could happen, for example, if all pLo and oLx data points at a given level of pLx fall in the same (or very small) area, implying no (or little) variance on either dimension. While this state of affairs might be interpreted as indicating a high degree of cognitive balance, it is more parsimoniously viewed as a manifestation of a simple response set. Such an effect, that is, may or may not be determined by tendencies toward cognitive balance. In contrast, a correlation (literally understood as covariation) cannot be explained on the basis of simple response stereotypy and is therefore better evidence for the operation of a higher order set implied by the term "cognitive balance."

measuring instruments. For example, testing the prediction of the theory regarding the elevation of the curve (value of k) in Equation 1, requires at least that the oLx scale have a true zero. Likewise, testing the theory with respect to the slope of the curve (value of w) in Equation 1, requires that the oLx scale have equal intervals and that its end points define maximum degrees of liking and disliking. Whether these assumptions are met by scales currently available is discussed more fully in an ensuing section.

Another way of testing the fit of the model to empirical data does not depend upon determining degrees of correlation. A more detailed statement of assumptions regarding the properties of the "tetrahedron model" would make it possible to construct a procedure for indexing the degree of balance of each possible point in the three-dimensional space. Such an index of the degree of balance of a structure could be given a relative frequency (likelihood) interpretation as suggested by Cartwright and Harary (1956), an interpretation in accordance with the assumption that imbalanced points are avoided and thus occur less frequently than balanced ones. If such a formula were constructed, and a set of operations carried out to associate a number with each possible point (or with each of a finite set of subspaces in the cube), the model could be tested by noting the amount of correspondence between theoretical and observed frequencies in the three-dimensional subspaces. It should be noted, however, that given the model alone, there does not exist a procedure for determining the theoretical frequencies in the subspaces. The model itself specifies nothing about the frequency distributions of the three constituent dimensions, pLo, pLx, and oLx, considered separately: That is, the model does not predict how much p likes o unless the relations between p and x and between o and x are given. Therefore, the theoretical frequencies are contingent upon the shape of the joint (bivariate) distributions of at least two of the variables. Given one of the three possible bivariate distributions, a frequency "expected" on the basis of the quantitative balance model could be assigned to each subspace in the cube.

This particular method of testing the model would provide the same information about the predictive utility of the model as that given by the earlier presented method based on covariation of any two variables at different levels of the third. The latter method based on covariation appears to be more easily carried out in practice and also embodies the possibility of stating the model as a set of linear equations (or monotonic functions), modes of representation that provide a readily understandable description of the fit of the data to the model.

Some aspects of the model just described may be compared with the theory of attitudinal congruity proposed by Osgood and Tannenbaum (1955). These theorists have introduced quantification into their model; however, they have quantified only two of the three relations in the three-element configuration. In Osgood's (1960) system of representing a simple cognitive structure, the signs attached to o and x represent a person's attitude toward the person or object, o and x, as measured on

the evaluative scales of the Semantic Differential (Osgood, Suci, & Tannenbaum, 1957). These evaluative measurements can vary from some negative number to some positive number. The sign of the line between o and x indicates whether the two elements are perceived to be closely related to one another or in opposition (or in the special case they discuss, whether a source, o, is believed to favor or to oppose an idea or object, x). The latter relation itself, however, is treated as a simple dichotomy; it is either positive (associative) or negative (dissociative). On this basis Osgood and Tannenbaum propose that attitudes toward objects are congruent (*a*) if they are equally polarized in the same evaluative direction when the objects are related by positive assertions or (*b*) if they are equally polarized in opposite evaluative directions when they are perceived to be in opposition.

How would these definitions of congruity be affected by varying the strength of the relation between o and x? A plausible answer is that decreasing the strength either of a positive or negative assertion relating o and x could only make the definition of congruity less certain or more indeterminate. Such a conclusion is suggested by the tetrahedron model. Perhaps Osgood and Tannenbaum have deliberately chosen to restrict consideration to the special case where at least one relation is of maximum strength to avoid dealing with cases in which predictions are less determinate.

A suggestion from the work of Osgood and Tannenbaum may also be relevant to the tetrahedron model. Within the context of Osgood's (1952) theory of meaning, it is concluded that there is a tendency for attitudes toward (evaluations of) objects to become polarized: That is, there is a continuing pressure toward polarization since extreme, "all or nothing," "black and white" judgments are simpler than more refined ones. Applying this reasoning to the tetrahedron model, one might hypothesize forces directed away from the center of the cube and toward its sides. This hypothesis implies that attitudes or cognitions that are unrelated to one another (forming vacuously balanced structures) should be relatively rare, a phenomenon also implied by Harary's (1959) postulated "tendency toward completeness." While forces directed away from the center of the cube would tend to make cognitive elements relevant to one another, it is the tendency toward cognitive balance that determines the form of the relations among these elements. In short, two classes of forces might be postulated, one tending to keep points way from the center of the cube (Osgood & Tannenbaum's pressure toward polarization) and the other (tendencies toward balance) tending to limit the distribution of points to certain (balanced) positions along the outer surface of the cube.

Application of the Model to Two Behavior Domains

While the quantitative version of Heider's balance theory is applicable to a wide variety of social psychological problems, the empirical studies reported here constitute essential first steps in providing direct and easily interpretable tests of the model.

The quantitative balance model was tested in two related behavior domains: (*a*) perceived interpersonal attraction among variously evaluated peers (Study I) and (*b*) perceived reciprocation of liking by others toward the self under various levels of self-esteem (Study II).

Study I tested the model in the context of p-o-q configurations (where another person, q, is substituted for x in Heider's p-o-x triad). The study has features in common with previous research (e.g., Kogan & Tagiuri, 1958; Tagiuri, 1958) on the perception of interpersonal preference, with the addition of variable strengths of L relations. The specific hypothesis tested is that there is a gradient in the degree of correlation between a person's own evaluative ratings of a set of others and the liking he perceives each of these others to feel toward (or to "receive from") various focal persons—the gradient being a function of the person's attraction to the focal person. The hypothesis can be stated as a simple monotonic function, $r_{\mathrm{pLo}\,,\ \mathrm{oLq}} = f(\mathrm{pLq})$, or as a linear equation, $r_{\mathrm{pLo}\,,\ \mathrm{oLq}} = k + w(\mathrm{pLq})$, where k and w are defined as for Equations 1, 2, and 3 in a preceding section. In other words, the extent to which a person's (p's) liking for a set of others (o's) is correlated with the amount of liking he perceives these others (o's) to feel toward some focal person (q), depends upon how much he (p) likes the focal person (q).

Study II tested the model for p-o-s configurations where s, the perceiver's self, is considered an element in the cognitive structure. Thus, Study II examined Heider's (1958, p. 210) assertion that his definitions of balanced and imbalanced structures require the assumption that p likes s (himself): That is, that the perceiver has high self-esteem. Research by Broxton (1963), Deutsch and Solomon (1959), Lundy (1956), Pilisuk (1962), and Secord, Backman, and Eachus (1964) lends support to Heider's suggestion that a person's attitude toward himself is highly relevant in the context of balance theory.

In terms of Heider's theory one can describe four balanced and four imbalanced p-o-s triads, analogous to the eight structures listed in Figure 1. The structure corresponding to Case I in Figure 1 is interpreted as follows: If p has high self-esteem, he should like those whom he sees as liking himself. Case III means, if p likes himself he tends to dislike o if he thinks he is disliked by o. If p genuinely dislikes himself he should like those whom he sees as disliking him (Case II) and dislike those he thinks like him (Case IV).

Whether these generalizations are valid, of course, can be determined only if there are persons who truly meet the condition, p dislikes s. It is likely, however, that currently measured low self-esteem implies a great deal of ambivalence about the self rather than genuine dislike. Also, the specific relation, p dislikes s, is rare or at least difficult to conceive, in part *because* this relation is imbalanced. A structure with the relation, p dislikes s, is always imbalanced if it is assumed that the unit (belonging) relation between p and s is positive.

However, a test of the general hypothesized relationship between p's degree of liking for s and the correlation (correspondence) between pLo and oLs—a relation derived from the quantitative extension of Heider's theory—does not require persons who genuinely dislike themselves; the only requirement is to be able to identify reliably different degrees of pLs.

In spite of the fact that no fully satisfactory conceptualizations of self-esteem are available, nor completely adequate measuring instruments (Wylie, 1961), Study II used a self-esteem scale to measure the dimension, pLs. Although it is not feasible to locate an indifference point on a scale of self-esteem nor even to expect that persons scoring low in self-esteem genuinely dislike themselves, it is reasonable to assume that low scorers have less net positive feelings about themselves than do high scorers. Thus, the specific hypothesis of Study II is that the degree of correlation between a person's liking for various others and his perception of how much they like him, varies positively with the person's level of self-esteem. This prediction can be represented as the monotonic function, $r_{\text{pLo}, \text{oLs}} = f(\text{pLs})$, or as the linear equation, $r_{\text{pLo}, \text{oLs}} = k + w(\text{pLs})$, where k and w are defined as in an earlier section.[3]

Method

Both Study I and Study II were conducted with subjects (*Ss*)[4] in elementary and junior high

[3] Several general statements identify issues that intentionally are not dealt with in these studies. First, while it is possible to examine the actual network of choices (liking relations) among individuals to discover the degree of balance of such social structures (Cartwright & Harary, 1956; Newcomb, 1953, 1961), this is clearly a different problem from that of ascertaining the degree of balance of these structures as perceived by individuals. Study I and Study II concentrate on the *perceived* or *cognized* relations among persons. Second, it is possible to view the degree of similarity between perceived and actual social structures as defining the "accuracy of social perception." But as Tagiuri, Blake, and Bruner (1953) and Cronbach (1955, 1958) have noted, the interpretation of such derived measures of accuracy is fraught with great difficulty. In short, the two issues, degree of balance of the actual social system and veridicality in social perception are not treated further here. Third, neither study attempts to explicate causal relations among the behavioral measures studied. The lack of concern with problems of cause and effect in the data is not a matter of inadvertence. Balance theory implies only that cognitive structures are stable systems in which a change in any part may produce changes in other parts. A test of the theory does not require a demonstration that any particular causal sequence is typical. For example, balance theory is noncommittal on the question whether perceived agreement typically leads to liking or whether it is more typical for liking to give rise to perceived agreement. Experimental studies (Backman & Secord, 1959; also see Rosenberg, 1960, in a more general context) support the view that *both* processes can occur. The present studies examine the degree of balance in naturally occurring, nonmanipulated cognitive structures, and therefore have no bearing on the priority of liking versus perceived agreement. Finally, the current studies bypass questions of the causal origin of the tendency toward balance. The author believes, however, that it may be profitable to examine the importance of such antecedent variables as social reinforcement history, variety of experience, and the current contingencies of reinforcement in the person's social group.

[4] A distinction is made between *S*, the conventional symbol for subject, and s, a theoretical term designating the perceiver's self. The fact that *Ss* are elsewhere referred to as p's merely reflects the use of Heider's term, p, for person or perceiver.

school, the data consisting of self-esteem scores and sociometric-like ratings of class members of the same sex. Each *S* was asked to indicate his degree of attraction for every other person in his group, his perception of the feelings of these persons toward himself, and his perception of the liking relations among selected other classmates. The latter ratings were obtained via two modes, each *S* responding under one mode only. Mode 1 required judgments about the attitudes of three focal persons (someone p likes, feels indifferent toward, and dislikes) toward all of the other classmates. Mode 2 reversed the task, requiring judgments of how everyone else in the class feels toward the three focal persons—a well-liked, a neutral, and a disliked other. Within each sex and grade level approximately one half of the *S*s performed under Mode 1 and one half under Mode 2.

Each *S*'s self-esteem was assessed both by a self-descriptive instrument and by teacher's ratings of certain behaviors.

Subjects. A total of 415 *S*s (197 boys and 218 girls) from 14 classrooms was drawn from Grades five, six, and seven. The classes were obtained from five different public schools, representing considerable variation in socioeconomic status. Each classroom contributed approximately 30 *S*s; however, because the average enrollment in each class was 34 (range 30–38) the average boy or girl rated approximately 16 classmates (the average number of persons of the same sex in the group, 17, minus 1, the child himself).

Variables and Instruments. Four separate sets of responses were taken in Parts A, B, C, and D of a questionnaire. Part A, designed to measure each *S*'s evaluation of all other classmates—pLo[5]—confronted each child with a list of the names of all classmates of the same sex and a 7-point rating scale with two extreme end points defined as "like very much" and "dislike very much." The midpoint of the scale was defined as "no feeling" and "neither like nor dislike." The other 4 points on the scale were also verbally defined.

Part B, a measure of each *S*'s perception of the degree to which he is liked by each relevant classmate—oLs—consisted of the same list of names and a similar rating scale phrased to evoke the appropriate judgments.

Part C, a 50-item Self-Esteem Inventory (SEI: Coopersmith, 1959) provided one measure of pLs, the degree of liking for the self. A second index of self-esteem was a 10-item Behavior Rating Form (BRF: Coopersmith, 1959). These two indexes, one based on self-rating and the second on teacher's ratings, were thought to have potential value when used in conjunction to distinguish "defensively" high scorers from those with genuinely high self-esteem. The former, according to Coopersmith (1959, p. 87), should describe themselves in quite favorable terms while their observable behavior would be indicative of low self-esteem.

Part D, a measure of oLq, the degree of perceived attraction among various other classmates, was composed of the same list of classmates and a rating scale phrased to evoke the rater's judgment of the liking relation among various other persons, o's and q's. To restrict the number of judgments required of *S*, three *focal persons* for each *S* were selected by the experimenter (*E*) to represent widely disparate values on pLo. Which three classmates were selected as focal persons for a given *S* depended only on that *S*'s pattern of ratings; thus, the focal persons were not the same individuals for all *S*s. (This procedure resulted in the average child's having to make 3×15, or 45, judgments rather than the complete set of all possible judgments, 16×15, or 240.) An attempt was made to select as one focal person someone whom the child had rated +3 (like very much), someone to whom he had assigned a zero (neutral), and someone whom he had rated −3 (dislike very much). Because not all children used the entire range of the rating scale in Part A, *E* selected one classmate *S* had rated highest (*H*), one he had rated lowest (*L*), and one representing an intermediate (*M*) value in *S*'s distribution of evaluative ratings.[6]

The two modes of making oLq judgments in Part D depended upon whether the three focal persons were treated as o's (the perceived *source* of affect—from whom various degrees of liking are perceived to radiate) or as q's (the perceived *object* of liking—toward whom the affective ratings of others are perceived to converge). All judgments were made about how o feels toward q, and not vice versa. Thus, Mode 1 required each child, p, to judge how he thought each of the three selected focal persons, o_L, o_M, and o_H, felt toward the remaining class members, the q's. Mode 2 reversed the task for p; he was required to judge how he thought the remaining class members, o's, felt toward each selected focal person, q_L, q_M, and q_H.

Derived Measures and Analytic Procedures. Because the predictions obtained from the tetrahedron model refer to the structure of cognitions within an individual, four intraindividual correlation coefficients (Pearson *r*) were calculated by electronic computer for each *S* (three coefficients for Study I and one for Study II). Each coefficient, calculated to three digits, was transformed to Fisher's z_r (Edwards, 1950) for the statistical analysis.

The correlation coefficients obtained for Study I are as follows. For each *S*, the degree of correla-

[5] Because the distinction between o and q is relevant only for the ratings in Part D, the precise but cumbersome expression, pLo(q) representing a person's set of evaluative responses toward all others in his group, will be replaced by the simpler expression, pLo, unless the more explicit expression is required.

[6] The three focal persons for each child are labeled o_H, o_M, o_L (or q_H, q_M, q_L) respectively, for the classmate given a high, medium, and low rating by *S* on the evaluative scale of Part A.

tion was computed between the ratings he assigned to classmates in Part A and his three sets of ratings in Part D—those with respect to a person Low, Medium, and High on pLo. The three resulting coefficients are designated r_{AL}, r_{AM}, and r_{AH}, respectively.

For Study II, only one correlation coefficient was determined for each *S*. Ratings assigned in Part A, indicating degree of liking for each other classmate, were correlated with ratings in Part B in which *S* judged the degree of liking of each other person for him. The resulting coefficient, r_{AB}, should be positively related to self-esteem according to the quantitative balance model.

Results

Study I. A simple Subjects × Conditions analysis of variance (Edwards, 1950, pp. 264–302) was performed on the z_r transformed correlation coefficients, r_{AL}, r_{AM}, and r_{AH}, for the entire sample of 415 *S*s, yielding an *F* ratio of 190 with 2/414 *df*; $p < .0001$. Because the design involved "repeated measurements," the mean square for Subjects × Conditions interaction was the error term used to test the significance of the main effect for Conditions—levels of pLo(q) (Edwards, 1950; McNemar, 1955; Walker & Lev, 1953). The weak form of the hypothesis of Study I, that the mean *r* should be an increasing monotonic function of the level of pLo, is strongly supported by the overall analysis of variance as well as by analyses performed separately for each of the 28 groups. While the *F* ratios of five groups fail to attain significance at the .05 level, four of these have *p* values between .05 and .10 and one has a *p* value between .10 and .20. Of the remaining 23 groups showing a levels effect significant at less than the .05 level, 9 are well beyond the .001 level, 5 beyond the .005 level, 4 beyond the .01 level, and 4 beyond the .025 level. All but three groups show the predicted trend: $\bar{r}_{AL} < \bar{r}_{AM} < \bar{r}_{AH}$. Each of these involves a single reversal such that $\bar{r}_{AL}$ is slightly larger (from .007 to .032) than $\bar{r}_{AM}$.

A graphical representation of the highly significant total levels effect for all *S*s is shown in Figure 4, giving evidence of a nearly linear trend from $\bar{r}_{AL}$ to $\bar{r}_{AH}$. The points are plotted to represent (along the abscissa) the average *S*'s pLo scale values for the three selected focal persons. These average values, varying only slightly across groups and between sexes and modes, were −.96, .93, and 2.92. In other words, the average level of liking for the three focal persons o_L, o_M, and o_H, approximated a −1, +1, +3 pattern rather than the desired −3, 0, +3 pattern—the result of a tendency for most *S*s to give a greater number of positive than negative ratings. For comparison, the broken line in Figure 4 shows the results for a subset of *S*s (N = 56) whose ratings of the three focal others on the pLo scale conformed to the "ideal" pattern, −3, 0, and +3.

Very decisive support for the weak version of the hypothesis of Study I can be claimed. However, if the hypothesis is stated more stringently as a linear equation, it is apparent that the elevation of the curve is greater and the slope of the curve is less than the balance model alone would predict. These inferences concerning the fit of data to model are not strictly proper, however, unless one makes rather strong assumptions about the pLo scale—assumptions that are detailed in an ensuing section.

Comparison of the results for all combinations of sex and mode revealed a substantial difference between Mode 1 and Mode 2 in the magnitude of $\bar{r}$ ($t = 4.65$, $df = 413$, $p < .0001$) but almost identical results for the two sexes ($t = .23$, $df = 413$). Inspection

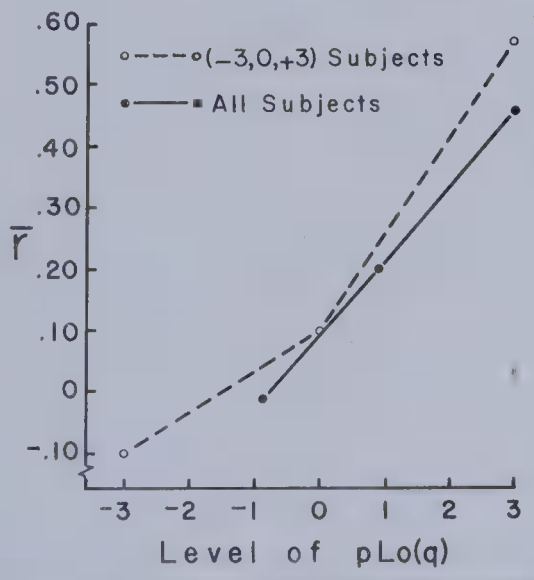

Fig. 4. Mean value of the dependent variable, *r* (determined by calculating mean z_r, then transforming to *r*) for all subjects (and for −3, 0, +3 subjects) at each of three levels of pLo(q).

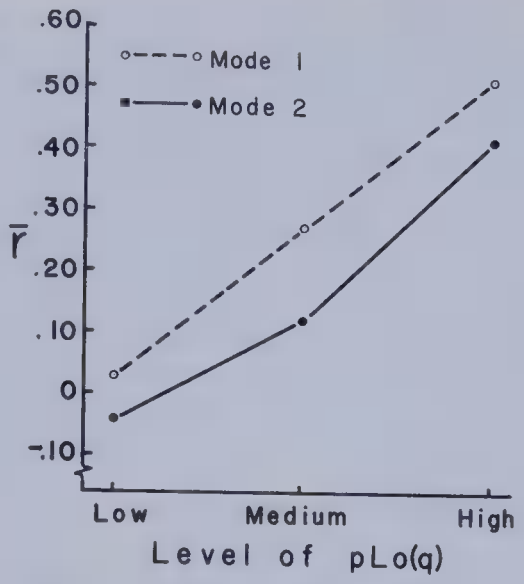

FIG. 5. Mean value of the dependent variable, r, for Mode 1 and Mode 2 at each of three levels of pLo(q).

revealed no slope differences between the sexes nor between the two modes. Figure 5 dramatizes the extent to which the curves for the two modes have a similar slope but a different elevation.

Study II. The hypothesis of Study II as derived from the extended balance model is that there is a positive correlation between *S*'s self-esteem and the measure of congruency, $z_{r_{AB}}$. The latter expresses the Fisher transformation of the correlation coefficient between *S*'s liking for a set of others (Part A ratings) and the extent to which he perceives his feelings to be reciprocated by these other persons (Part B ratings).

The two indexes of self-esteem, Coopersmith's (1959) SEI and BRF, are significantly ($p < .001$) related to each other (with some nonsignificant variation in this relation across groups). Because the correlation coefficient between the two scales is only approximately .40 and because both scales have considerably higher reliability coefficients than .40 (cf. Coopersmith, 1959), it appears that each scale measures a somewhat different aspect of self-esteem. Such an inference reinforced the view that the two scales have potential value as joint predictors of congruency. However, before employing SEI and BRF as joint predictors of congruency, their separate correlations with the congruency index are examined.

The degree of correlation between congruency and SEI was calculated for each of the 28 groups. The variation in these coefficients over the 28 groups was not sufficient to reject the hypothesis that the coefficients are from a single population. The mean of the correlation coefficients calculated separately within each group are extremely close to those obtained by calculating one coefficient for all boys, one for all girls, and finally one for the entire sample as a single group. Therefore, only the latter are reported in Table 1. All correlations between SEI and congruency are highly significant ($p < .0005$) for boys, girls, and all *S*s. Visual inspection of each scatter plot revealed no evidence of a curvilinear relation between SEI and congruency. Although the correlation between SEI and congruency is not of large magnitude, it is in the direction predicted by the extension of Heider's theory and is highly

TABLE 1

CORRELATION COEFFICIENTS BETWEEN CONGRUENCY[a] AND TWO MEASURES OF SELF-ESTEEM TAKEN SINGLY AND IN COMBINATION

Index of self-esteem	Boys		Girls		Both sexes	
	N[b]	r	N[b]	r	N[b]	r
SEI	196	.240	218	.222	414	.225
BRF	170	.245	186	.077	356	.131
SEI and BRF (Multiple R)	170	.292	186	.224	356	.234
SEI + BRF for $\lvert Z^{c}_{SEI} - Z_{BRF}\rvert \leq 5$	63	.244	67	.253	130	.226

[a] Congruency = $z_{r_{AB}}$ = Fisher transformation of Pearson correlation coefficient between pLo (Part A ratings) and oLs (Part B ratings).

[b] Unequal Ns derive from lack of BRF scores in two school classes and from nature of selection procedure for last index.

[c] $Z = 50 + 10z$ = standard score.

significant for the entire group ($t = 4.4$, $df = 414$, $p < .0005$).

Table 1 shows also that BRF is correlated significantly with congruency ($p = .0075$), but this relation is accounted for mostly by the boys. The sex difference in the correlation with BRF is highly significant with a two-tailed test ($p < .001$), the correlation between BRF and congruency for the boys being as high as that between SEI and congruency, whereas for the girls it is virtually zero. The three mean correlations obtained by combining (transformed) correlations for all boys' groups, for all girls' groups, and for all groups of both sexes, yielded estimates of r nearly identical with those given in Table 1.

Combining the two measures, SEI and BRF, in a multiple-regression equation yielded multiple Rs for boys, girls, and the entire sample barely distinguishable from the corresponding correlations between congruency and SEI alone. The relative weights or partial regression coefficients expressed in standard score form (Walker & Lev, 1953, p. 319) of SEI and BRF in predicting congruency reflect the already noted sex difference: That is, for boys the two coefficients are nearly equal (.167 and .180 for SEI and BRF, respectively) while for girls, the two coefficients are quite different (.230 and −.033 for SEI and BRF, respectively).

Another method of combining the two indexes to predict congruency involved summing the two standard scores for those *S*s whose SEI and BRF scores were not more than one-half standard deviation apart, a method closely resembling one recommended by Coopersmith (1959). The method assumes that *S*s whose SEI and BRF scores are in agreement best represent different positions on a dimension of genuine self-esteem: That is, such a selection procedure should exclude both the "defensively" high scorers as well as those whose habitual verbal self-descriptions reflect less favorable conceptions of themselves than their observable school behavior indicates.

For the 130 *S*s whose self-descriptions agreed with the teacher's ratings of them, the correlation between congruency and ($Z_{SEI} + Z_{BRF}$) is .226, a significant value ($p < .005$) but one involving no improvement in prediction of congruency over that provided by SEI alone. The scatter plot was inspected visually on the chance that there might be an unexpected curvilinear relation between congruency and this "refined" index of self-esteem. No such relation appeared.

Discussion

Study I

For Study I the model's prediction was stated both as an increasing monotonic function and as a linear equation, the latter statement requiring "stronger" assumptions about the properties of the measuring scales. These two ways of representing the model are designated the "weak" and the "strong" versions, respectively.

The Balance Theory As a Monotonic Function. The results of Study I decisively support the weak version of the extension of Heider's theory. The expected trend in the magnitude of correlation coefficients across different levels of pLo is an outstanding feature of the results for all groups combined (as well as being found rather consistently within the groups considered separately) and the statistical significance of the results leaves little doubt about the reliability of this finding. In concrete terms, the more a person, p, likes someone, o(q), the more positive is the correlation between p's evaluation of a set of other persons and p's perception of o's feelings toward the same set of other persons (or p's perception of the feelings of this set of others toward q). In other words, a person's cognitions of the relations among others (when these relations are measured on at least an ordinal scale) tend toward the balanced or consistent organization implied by the quantitative model.

The Balance Theory As a Linear Equation. The extent to which the results accord with the strong version of the model is more difficult to determine; it depends not only upon the obtained results but also upon whether one can assume that the pLo scale has at least the following properties: a zero representing a true affective indifference point and polar extremes (+3 and −3) representing maximum degrees of lik-

ing and disliking, respectively. On the face of it, the pLo scale is compatible with these requirements. That is, every position on the 7-point scale is given a straightforward and explicit verbal definition in terms as unambiguous as possible.

The fact that the curve in Figure 4 for all *S*s appears to be linear and exhibits a slope significantly greater than zero can be interpreted as confirmation of the model's predictions. But if one applies a more rigorous test, the fact that the slope is not so great as it is predicted to be can be taken as evidence against the model. That is, if the quantitative balance model is assumed to predict that all cognitive structures are balanced,[7] then these data might be taken to provide another "exception to the rule" that Zajonc (1960) suggests may become increasingly characteristic of research findings on consistency theory. In his comparison of attitudinal consistency theories ("human nature avoids inconsistency, imbalance, or dissonance") with the earlier concept of vacuum in physical science ("nature abhors a vacuum"), Zajonc (1960) notes that in both cases the principle systematically accounts for many phenomena but that there are too many exceptions to consider it a theoretically useful generalization. In the present instance, however, it is not completely clear that the discrepancy between predicted and obtained slope of the curve in Figure 4 should be interpreted as an "exception" to the prediction of balance theory. It is quite possible, for example, that the results obtained are precisely those to be predicted by balance theory if information were available about some other variable in the situation. Such a possibility does not imply that the quantitative balance theory is so flexible that it can apparently "account for" any data, no matter what their pattern. Instead, the theory suggests the kinds of variables that should be examined for their possible relevance in improving the fit of any given set of data to the model. For example, it can easily be shown that the correlational data presented in Figure 4 would more closely accord with the model's predictions *if* it were known that the data had been obtained under a special condition characterized by a less than maximal value on some fourth relevant variable. More specifically, it is possible that among the young *S*s studied (ages 11–15), peers of the same sex are simply not very important or salient as objects of expressed and perceived affect. Indeed, if one adds to the three variables in the structure, a fourth, *perceived importance* of interpersonal relations, it can be shown that the predictions of the tetrahedron model are only special cases that assume maximally high values on a scale of perceived importance (as well as on all other relevant variables). If perceived importance is minimal (zero), the predicted slope of the curve would be zero rather than +1. In general, the higher the perceived importance, the closer would one expect the data to approximate the

[7] The reader is reminded that in the context of the present three-dimensional quantitative model, any cognitive structure lying within the tetrahedron of Figure 2 is defined as "balanced." This definition, of course, calls many more structures balanced than would a strict adherence to Heider's rule as it is interpreted in Figure 2. To the extent that balance, in this sense, characterizes *S*'s entire set of measured triadic structures, the data would fit the equations given earlier. Any deviation from the predicted slope of the line (cf. equations on p. 4) could indicate that at least some cognitive structures are imbalanced. Specification of which particular triadic structures are imbalanced and to what degree would require separate examination of each structure. If one wishes to treat the entire set of cognitions and/or feelings expressed by *S* in a given content area as a single cognitive structure, then the slope of the regression line (the value of w in equation 3, p. 4) or the coefficient of correlation between $r_{pLo,\ oLq}$ and pLq can be interpreted as an index of the degree of balance of *S*'s cognitions in that domain.

From a somewhat different perspective, one might also regard this correlation (balance) coefficient as an index of the reliability (internal consistency and precision) of a speaker's verbal behavior. Such a balance coefficient might be useful to the psycholinguist in describing the course of a child's acquisition of language or in describing differences among various language groups. Presumably, only a trained logician, taking pains to edit his verbal behavior in accordance with the rules of formal logic could obtain a balance coefficient of +1.00. A child learning to speak might display a balance coefficient of only slightly greater than zero, while a schizophrenic exhibiting "opposite speech" might have a negative coefficient.

predictions of the simple three-dimensional tetrahedron model.

There is probably no reason why the quantitative balance model cannot be extended to accommodate any finite number of variables. Perceived importance is offered here as only one of several possible explanatory variables; self-esteem is another variable that might act as a moderator variable in the same way. Whether incorporating these variables or any others into the model actually improves the correspondence between data and theory must, of course, be determined empirically. The question would best be answered by direct experimental manipulation of those variables assumed to be relevant.

The possibility that measurement error is one principal source of the less-than-predicted slope of the curves in Figure 4 should also not be overlooked. In other words, the evaluative ratings may be inherently unreliable or may be unstable over time; the possibility that the depressed correlations resulted from reliable fluctuations over time in degree of liking is especially noteworthy since up to 7 days elapsed between the ratings of Parts A, B, and C and those of Part D. To what extent the lack of perfect cognitive consistency implied by the obtained slope of the curves should be interpreted as the result of some irreducible measurement error or as the result of failure to consider other identifiable relevant variables in the structure, remains an important problem for future research.

A quite different way to account for the smaller-than-expected slope of the obtained curve invokes possible weaknesses (other than unreliability) of the pLo scale itself. For example, it may be that the scale does not discriminate finely enough nor extend far enough on its positive end—a criticism less likely relevant for the negative end of the scale because few extremely negative ratings occurred. That most *Ss* manifested a favorability response set in rating their classmates would tend to reduce the scale's effectiveness in discriminating extremely well-liked from moderately well-liked classmates. Those *Ss* who did distribute their ratings over the entire scale apparently reserved the highly positive end of the scale for those they truly like very much. Such an interpretation is supported by the fact that $\bar{r}_{AH}$ is considerably higher for the -3, 0 $+3$ *Ss* than for the total sample at the "same" highly positive value of pLo(q) (cf. Figure 4).

A second feature of the data of Study I is the greater elevation (larger value of k) of the curve than the model predicts. The upward elevation of the curve means that, given the slope of the line, the correlations are more positive than expected.

The fact that $\bar{r}_{AL}$ is near zero rather than highly negative indicates that persons attribute merely different rather than opposite evaluations to class members they dislike most (or like least). Only very slightly opposing (negatively correlated) evaluations are attributed to classmates disliked very much (i.e., assigned a value of -3 on the pLo scale). Also, whereas the model predicts that the feelings of an "indifferent" person toward others (or toward an "indifferent" person by others) are perceived to be unrelated to a person's own feelings toward these others, they in fact are perceived to be positively related to his own evaluations of others. Even in those instances in which a classmate explicitly is assigned a zero (indifference) rating on the pLo scale, $\bar{r}_{AM} = +.10$.

Harary (1959) has described the above-noted phenomenon as a "tendency toward positivity." Rosenberg and Abelson's (1960) concept of a "force to maximize potential gain and minimize potential loss" as well as what McGuire (1960) has labeled "wishful thinking" are also analogous tendencies that sometimes are opposed to balance and logical thinking, respectively.

Perhaps the "positivity effect," as reflected in the correlations, is a special case of the general tendency to assume similarity with others (Cronbach, 1955, 1958; Fiedler, 1958; Fiedler, Warrington, & Blaisdell, 1952). According to this view, the tendency to assume that others are similar to oneself carries over to situations in which one is trying to guess how a disliked person (or someone neither liked nor disliked) evaluates a set of others. This tend-

ency makes it difficult consistently to take the role of a disliked (and presumably dissimilar) other person.

An explanation of the heightened elevation of the curve is possible also in terms of what DeSoto (1961) has called a "predilection for single orderings." This account stems from the common sense idea that it is simpler to order objects on only one dimension than to order them in many ways. A greater amount of effort presumably is required to construct, to learn, or to remember a set of multiple orderings than a single ordering of a set of objects. In the specific case at hand, while *S* is attempting to order persons (actually, to rate them) as someone else would, his own ordering of the persons continually intrudes and presumably distorts the attempt to order from another person's point of view.

Another possible determinant of the greater-than-predicted elevation of the curve is that the zero point on the pLo scale may not truly represent psychological indifference. It has already been noted that *S*s do not tend to use the negative end of the scale and that the mean rating is very close to +1; this may be interpreted as a "favorability response set." Taken by itself this interpretation suggests that the psychological indifference point probably falls near +1 and that −3 represents a rather extreme degree of disliking. On the other hand, it is possible that the mean rating of +1 is not the product of a *mere* response set (favorability), but results instead from *S*s' genuinely positive feelings toward most peers. Such attraction would be of even greater magnitude than it appears to be if the evidence in Figure 4 is taken literally. In fact, if the truth of the balance theory is assumed, the evidence in Figure 4 indicates that the psychological indifference point falls close to −1, the pLo level at which $\bar{r}$ is zero, and that −3 represents only a small degree of disliking. These inferences can properly be checked only by further research with more highly refined scales. For example, a point of affective neutrality could be established on an improved pLo scale by using other information known to be associated with indifference; information on intensity, certainty, or confidence (Cantril, 1946; Katz, 1944; Stouffer, 1950; Suchman, 1950) or on response latency (Osgood et al., 1957) could be used to establish empirically a point or region of psychological indifference.

Interaction between Observations and Theory. In the two instances just discussed—evaluating the results with respect to both the slope and the elevation of the curve—two main alternative explanations are available. The explanatory dilemma is such that either one must assume the truth of balance theory as one way to establish certain properties of the measurement scales, or one must make certain assumptions about the measurement scales to determine whether and in what respects balance theory has been confirmed. This explanatory dilemma is, in principle, common to all scientific endeavor, and it is resolvable only by what some have called circular argument; for the careful investigator the result is involvement, not in a "vicious circle," but in a "spiral" of increasing confidence in both theory and measuring scales as each is successively refined in research. In the present instance, the strong version of the model should be checked more closely with measuring scales whose interval size and point of origin are more fully established.

The above dilemma does not arise when the weak version of the quantitative balance model is evaluated, because the required assumption that the pLo scale has at least the properties of an ordinal scale (Stevens, 1950) seems beyond controversy.

Mode Differences. Another unexpected but highly reliable finding of Study I is that the judgments obtained under Mode 1 are more highly correlated than are the "same" judgments under Mode 2. Detailed examination of the differences in the tasks required by the two modes yields at least two plausible ways to account for this fact.

If the judgmental task of Part D is conceptualized as a two-stage process that requires (*a*) adopting a set (taking the role of another person) and (*b*) making judgments within the context of the set, and if it is assumed that the first process requires more information (or energy)—adopting a

set implies preparing to make a variety of individual judgments—then it can be shown that Mode 2 requires not only a greater number of acts (changes in psychological state) but also a greater number of the difficult ones than does Mode 1. Under Mode 1, S was involved in $3N + 3$ acts (3 changes in set and from within each, N judgments); each change in set is an act and there are, in addition, N acts in each of 3 sets. If the average number of others for whom Part D judgments were made is 15 ($N = 15$) then Mode 1 required 48 acts. Under Mode 2, S performed $2 \times 3 \times N = 6N$ acts ($3N$ set adoptions and from within each set, a single judgment). The specific task under Mode 2 was to adopt a set, make a judgment within that set, adopt a second set, make another judgment, etc. until $3N$ sets had been adopted and a judgment made within the context of each. Thus if $N = 15$, Mode 2 required 90 separate acts—nearly twice the number required in Mode 1. Mode 2 also required 15 times as many of the presumably more demanding acts (adopting a set) than did Mode 1. If the more difficult task of Mode 2 results in less reliable judgments, the correlations for Mode 2 should be lower (closer to zero) than the correlations for Mode 1 subjects. Were it not for an overall "positivity effect" noted earlier, the presumed greater difficulty of the task under Mode 2 (and consequent lower reliability) also would result in a curve with lesser slope for Mode 2, since both negative and positive correlations would be closer to zero. Obtaining judgments under both Mode 1 and Mode 2 in a design that includes highly negative positions on the pLo scale would provide a test of this hypothesis: That is, the functions would cross if Mode 2 judgments were less reliable than those of Mode 1.

Another explanation of the Mode difference derives from the difference between perceiving a person as liking (disliking) and seeing him as being liked (disliked). The task of Mode 1 makes it easy for S to forget that he is to rate from another's point of view since, once he has adopted the set of another, he is not repeatedly reminded of this orientation while he judges how that person feels toward each other person in the series. In other words, as the series of judgments are made under Mode 1 the adopted orientation may become weaker (because it is not continually reinstated) and its place gradually taken by S's own pervasive orientation. In contrast, only one judgment at a time is made from within the context of a given set under Mode 2; in this case S judges how much a given person is liked by each of a set of others. Each judgment S makes necessarily occurs immediately after he has adopted the orientation of another person; hence S's own orientation is less likely to intrude as the primary framework within which the ratings are made. In short, the conditions of Mode 1 more easily allow S to lapse into assuming similarity (or into manifesting his own single ordering) than do the conditions of Mode 2. Since increasing either of these dispositions (assuming similarity with others or manifesting one's own ordering) would produce more highly positive correlations among S's ratings, one would predict Mode 1 correlations to be higher than those for Mode 2.

Study II

Only the weak version of the model was tested for Study II inasmuch as it seemed very unlikely that the pLs (self-esteem) scale possessed the required characteristics for a test of the strong version: That is, the results were evaluated only with respect to whether the obtained regression line has a slope greater than zero.

Although the relations among congruency and the measures of self-esteem are not strong enough to enable efficient prediction of individual differences, they do reach high levels of statistical significance. That the correlations are not higher can be attributed in part to relatively undeveloped and inadequately formulated instruments for assessing self-esteem.

Concretely, the findings of Study II mean that the higher a person's self-esteem, the more positive is the correlation between his feelings toward a set of others and his perception of their feelings toward him. While it is true that persons generally like those who they think like them and dislike those who they think dislike them, the main im-

port of Study II is that this generalization is most strongly manifested in persons with high self-esteem.

If balance theory is taken seriously, it is possible to suggest one reason for the relatively low correlations between self-esteem and congruency. The fact that the mean index of congruency is very high ($\bar{r}_{AB} = .74$) can be interpreted as evidence that, on the average, the sample *S*s hold quite positive conceptions of themselves. The rather marked restriction on the range of self-esteem, compared to the range of feelings toward other objects or persons, could be an important determinant of the low degree of measured correlation between self-esteem and congruency.

The relation between self-esteem and congruency was found for both measures of self-esteem—SEI and BRF. However, the correlation with the latter index (that based upon observer's ratings) varies with sex. Apparently the behavior of a fifth-, sixth-, or seventh-grade boy, as it is observed by a teacher, is somewhat related to his feelings about himself and his perception of his relations with peers, whereas this does not hold for the sample of girls.

Several related and rather general theoretical implications of the two studies warrant further comment. The first is the problem of defining the boundary conditions within which cognitive balance can be expected. A second problem concerns individual differences in what some authors call "tolerance for imbalance." A third involves the question of multiple routes to achieving balance (alternative methods of dissonance reduction) in the context of experimental studies of attitude change.

The findings of both Study I and Study II clearly support one of the most important theoretical implications of the quantitative balance model—that the tendency toward balance is strongest when all perceived and/or felt relations are maximally intense. For example, the correlation between pLo and oLq is maximally high only when p feels relatively strongly (intensely) about q. This is shown directly for the case of intense positive feelings by the data of Figure 4. Such a conclusion is also supported by a further study by the author (in preparation) as well as one by Kanouse (1964). The latter studies show that *very* high levels of attitudinal consistency are exhibited in attitudes toward controversial social issues and public personalities evoking intense feelings. For example, those extremely favorable toward Albert Schweitzer judge Schweitzer's position as nearly identical to theirs on a wide variety of issues including abstract art, capital punishment, comic books, legalized gambling, movie censorship, racial segregation, and the United Nations. Conversely, persons extremely unfavorable toward Fidel Castro, for example, tend to judge his stands as nearly opposite theirs on the same set of issues. Positive correlations between *S*'s own views and the assumed views of an admired person and negative correlations with the assumed views of the villain reach average values of .80 and −.30, respectively. Perhaps these correlations could be made even higher with an effective experimental manipulation of the intensity of *S*'s attitude toward the persons or issues in question.

The fact that the correlations are not typically +1.0 or −1.0 suggests that complete balance (i.e., balance among *all* elements of an attitude cluster) is a very unusual state requiring special conditions. Exactly what these special conditions are is not yet clear, but the present formulation suggests that maximally polarized values (maximal distance from a zero or indifferent point) on all variables in the structure are required. Because these rather special conditions are rare in naturally occurring attitude structures (and may be difficult to produce experimentally) balance theorists have been justifiably cautious in speaking of balance as a "tendency" rather than a state that is unequivocably characteristic of an attitude cluster. The strength of the quantitative balance model in the face of these considerations is that it defines just those (perhaps unrealizable) conditions under which complete balance would be found and, more important empirically, it predicts the degree of balance that will occur under any values of the relevant constituent variables (relations). Note that under some circumstances the prediction made by this model may be

simply, "there will be no correlation among the ratings." While predicting the null hypothesis is generally uninteresting (and even undesirable), it becomes of considerably greater interest when it is placed in the context of a more general set of predictions (Suppes, 1964). In Study I, for example, such a prediction is simply a special case of the general linear equation; the special case can be stated, "if pLo is zero then there will be no correlation between pLq and oLq." Another example in which predicting the null hypothesis is a special case of a more general set of predictions might be: "if *S*'s self-esteem is extremely low, there will be no relation between $r_{pLq\ ,\ oLq}$ and pLo" (cf. Figure 4). That is, the line of regression will not have a slope greater than zero.

The findings of both Study I and Study II also may have implications for the problem of specifying the basis of individual differences in what has been called "tolerance for inconsistency," or "tolerance for dissonance" (Bem, 1961[8]; Feather, 1964; Festinger, 1957; Hovland & Rosenberg, 1960). It might be suggested that any variable reliably correlated with individual differences in "tolerance for imbalance or inconsistency" should be conceptualized as simply another relevant variable in the structure moderating the relationships among other response variables. This treatment places such additional variables systematically within the framework of balance theory. For example, Study II strongly supports Heider's (1958) conclusion that the usual illustrations of balance in the p-o-x triad require the assumption that p likes himself. Applying an analogous conclusion to the more complex structures of Study I suggests that *S*s with a high degree of "tolerance for inconsistency" (*S*s, that is, whose attitude structures are poorly predicted by a balance theory that ignores individual differences) may have generally low levels of self-esteem. It is likely, also, that such *S*s have idiosyncratic reinforcement histories with respect to the verbal behavior in question. If either of these possibilities were true, then treating individual differences in self-esteem (or degree of socialization) as a fourth variable in the structure would significantly increase the proportion of variance accounted for within the framework of balance theory.

Whatever their source, observed individual differences in tolerance for inconsistency constitute a problem warranting further study. Kanouse's (1964) attempt to relate individual differences in achieved balance to several personality measures, as well as some of the data on self-esteem in Study II constitute useful beginning ventures. Future research on this question should probably consider the trait dimension Newcomb (1963, p. 385) has discussed as "autism-realism." This trait may be relevant to individual differences in tolerance for imbalance in view of the likelihood that the perception of interpersonal relations (or the perception of others' attitudes in general) may reflect a compromise between complete cognitive consistency and the constraints of reality. According to this view, one with less than perfect balance is likely to be more "realistic" than one who exhibits completely balanced cognitive structures.

Finally, the proposed quantitative model has implications for the well-known observation (Heider, 1958, pp. 207–209) that there are several ways to eliminate imbalance or inconsistency, once it has been produced. This issue is of greatest interest in the context of studies which attempt to produce attitude change by direct experimental manipulation of variables. For example, Rosenberg and Abelson (1960, p. 121) following in part an earlier paper of Abelson's (1959), discuss essentially three ways of redressing imbalance: (*a*) changing one or more signs, (*b*) redefining or differentiating concepts, and (*c*) ceasing to think about the matter. Festinger's (1957, p. 264) somewhat analogous set of three methods of reducing dissonance are: (*a*) changing the evaluation of one or more of the elements involved in dissonant relations, (*b*) adding new consonant cognitive elements, and (*c*) decreasing the importance of the elements involved in dissonant relations. The work of other investigators points to other routes

[8] Unpublished manuscript, 1961, titled "Is there a tolerance for dissonance?"

including, for example, changing degree of perceived choice (Cohen, 1960), "projecting" undesirable traits onto a person similar to oneself (Bramel, 1962), and being incredulous (Osgood & Tannenbaum, 1955) or not trusting the communicator (Hovland & Weiss, 1951). These are merely a few of the numerous possibilities; in fact, there should exist at least as many methods of reducing imbalance as there are relevant variables in the cognitive structure. In other words, each dimension in the quantitative balance model defines one variable that is a potential candidate for change when inconsistency is introduced into the system. Which variable will actually change or whether all variables will change slightly probably depends upon a number of factors, including the amount of effort required to effect change on a given variable (Rosenberg & Abelson, 1960, p. 133). In practice, investigators have usually tried to study change in one variable at a time, and they do so by attempting to block alternative routes to reducing imbalance. For example, to produce attitude change as a function of a persuasive communication, it is necessary to ensure that *S* perceives the communicator as trustworthy (Hovland & Weiss, 1951) and the issue as important (McGuire, 1960a, 1960b, 1960c), etc. Similarly, in a choice experiment, to produce increased attractiveness of a chosen alternative and decreased attractiveness of unchosen alternatives, it is necessary that *S* feel he had a "free choice" (Cohen, 1960) and that *S* not feel the whole matter is really unimportant (Festinger, 1957), etc.

At this juncture the quantitative balance model is useful in providing a rationale for the systematic interpretation of the variety of possible methods to reduce dissonance or eliminate imbalance. When more sophisticated measurement of these different variables and their resistance to change becomes available it should be possible to test more precise predictions. For example, given a pattern of experimenter-produced changes in one or more independent variables, the model should predict which variables are likely to change and how much change is expected on each.

SUMMARY

The investigation addressed the problem of measuring the degree of consistency or balance of cognitive structures for the general case in which the relations among elements in the structures are treated as continua rather than simple dichotomies. The quantitative extension of Heider's theory of cognitive balance offered here enables systematic treatment of cognitive structures whose elements have any degree of similarity or attraction to one another.

The extended balance model was tested in two behavior domains: perceived interpersonal attraction among variously evaluated peers (Study I) and perceived reciprocation of liking by others toward the self under various levels of self-esteem (Study II). Both studies were conducted with 14 groups of boys and 14 groups of girls in School Grades five, six, and seven. Each of the 415 children in the sample provided the following data: self-ratings on a self-esteem inventory, a series of sociometric-like judgments indicating degree of liking for all classmates, perception of the extent to which he is liked by each peer, and perception of the liking relations among selected classmates. The last named judgments were obtained via two modes with a given *S* performing under only one mode. Mode 1 required judgments about how three focal persons (a well-liked, a neutral, and a disliked classmate) felt toward the remaining classmates, and Mode 2 required judgments about how each other person in the class felt toward the three focal persons. Teacher's ratings of pupil behavior related to self-esteem were also obtained.

Implications regarding intercorrelations among particular sets of *S*'s judgments were deduced from the quantified version of Heider's theory. In Study I the theory predicted a gradient in the degree of correlation between *S*'s own evaluative ratings of a set of others and the ratings he perceived radiating from or converging upon the three focal persons, the gradient being a function of *S*'s attraction to each focal person. For Study II the theory predicted that the degree of correspondence between *S*'s

liking for various others and his perception of how much they liked him, varied positively with *S*'s level of self-esteem. These predictions were confirmed at very high levels of statistical significance for both studies.

Study I revealed that the perceived preference relations among others are characterized by a high degree of balance as predicted by the quantitative version of Heider's theory but that the degree of balance is less than perfect; this result is discussed as possibly reflecting the moderating effects of other relevant variables. Correlations among sets of judgments were also generally more positive than predicted by the theory—a result consistent with the presumed operation of a "positivity tendency." No difference between the sexes appeared, but a striking difference was found between the judgmental modes; the intra-individual correlations were more positive in Mode 1 than in Mode 2, a finding discussed in terms of certain structural differences between the two modes.

Study II demonstrates that the extent to which a person perceives that his feelings toward others are reciprocated varies positively with his level of self-esteem. This relation is the same for both sexes when a self-rating device is used to assess self-esteem.

Implications of the model and the findings for several kinds of research problems are discussed. The model suggests new and more highly differentiated hypotheses, and it contributes to the possibility of developing better measuring scales. It is concluded that the quantitative extension of Heider's theory and the method proposed for testing the theory are effective media in which to study properties of cognitive balance.

REFERENCES

Abelson, R. P. Modes of resolution of belief dilemmas. *Journal of Conflict Resolution,* 1959, **3,** 343–352.

Abelson, R. P., & Rosenberg, M. J. Symbolic psychologic: A model of attitudinal cognition. *Behavioral Science,* 1958, **3,** 1–13.

Backman, C. W., & Secord, P. F. The effect of perceived liking on interpersonal attraction. *Human Relations,* 1959, **12,** 379–384.

Bramel, D. A dissonance theory approach to defensive projection. *Journal of Abnormal and Social Psychology,* 1962, **64,** 121–129.

Brehm, J. W., & Cohen, A. R. *Explorations in cognitive dissonance.* New York: Wiley, 1962.

Broxton, June A. A test of interpersonal attraction predictions derived from balance theory. *Journal of Abnormal and Social Psychology,* 1963, **66,** 394–397.

Burdick, H. A., & Burnes, A. J. A test of "strain toward symmetry" theories. *Journal of Abnormal and Social Psychology,* 1958, **57,** 367–370.

Cantril, H. The intensity of an attitude. *Journal of Abnormal and Social Psychology,* 1946, **41,** 129–135.

Cartwright, D., & Harary, F. Structural balance: A generalization of Heider's theory. *Psychological Review,* 1956, **63,** 273–293.

Cohen, A. R. Attitudinal consequences of induced discrepancies between cognitions and behavior. *Public Opinion Quarterly,* 1960, **24,** 297–318.

Coopersmith, S. A method for determining types of self-esteem. *Journal of Abnormal and Social Psychology,* 1959, **59,** 87–94.

Cronbach, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin,* 1955, **52,** 177–194.

Cronbach, L. J. Proposals leading to analytic treatment of social perception scores. In R. Tagiuri and L. Petrullo (Eds.), *Person perception and interpersonal behavior.* Stanford, Calif.: Stanford Univer. Press, 1958. Pp. 353–379.

DeSoto, C. B. The predilection for single orderings. *Journal of Abnormal and Social Psychology,* 1961, **62,** 16–23.

Deutsch, M., & Solomon, L. Reactions to evaluations by others as influenced by self-evaluations. *Sociometry,* 1959, **22,** 93–112.

Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.

Feather, N. T. Acceptance and rejection of arguments in relation to attitude strength, critical ability, and intolerance of inconsistency. *Journal of Abnormal and Social Psychology,* 1964, **69,** 127–136.

Festinger, L. *A theory of cognitive dissonance.* Evanston: Row Peterson, 1957.

Fiedler, F. E. Interpersonal perception and group effectiveness. In R. Tagiuri and L. Petrullo (Eds.), *Person perception and interpersonal behavior.* Stanford, Calif.: Stanford Univer. Press, 1958. Pp. 243–257.

Fiedler, F. E., Warrington, W. G., & Blaisdell, F. J. Unconscious attitudes as correlates of sociometric choice in a social group. *Journal of Abnormal and Social Psychology,* 1952, **47,** 790–796.

Harary, F. On the measurement of structural balance. *Behavioral Science,* 1959, **4,** 316–323.

Heider, F. Attitudes and cognitive organization. *Journal of Psychology,* 1946, **21,** 107–112.

HEIDER, F. *The psychology of interpersonal relations.* New York: Wiley, 1958.

HOROWITZ, M., LYONS, J., & PERLMUTTER, H. Induction of forces in discussion groups. *Human Relations,* 1951, **4,** 57–76.

HOVLAND, C. I., & ROSENBERG, M. J. Summary and further theoretical issues. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm, *Attitude organization and change.* New Haven: Yale Univer. Press, 1960.

HOVLAND, C. I., & WEISS, W. The influence of source credibility on communication effectiveness. *Public Opinion Quarterly,* 1951, **15,** 635–650.

JORDAN, N. Behavioral forces that are a function of attitudes and of cognitive organization. *Human Relations,* 1953, **6,** 273–287.

KANOUSE, D. *A quantified cognitive consistency model applied to social attitudes.* Unpublished BA thesis, Reed College, 1964.

KATZ, D. The measurement of intensity. In H. Cantril (Ed.), *Gauging public opinion.* Princeton, N. J.: Princeton Univer. Press, 1944. Pp. 51–65.

KOGAN, N., & TAGIURI, R. Interpersonal preference and cognitive organization. *Journal of Abnormal and Social Psychology,* 1958, **56,** 113–116.

LUNDY, R. M. Self perceptions and descriptions of opposite sex sociometric choices. *Sociometry,* 1956, **19,** 272–277.

MCGUIRE, W. J. A syllogistic analysis of cognitive relationships. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm, *Attitude organization and change.* New Haven: Yale University Press, 1960. (a)

MCGUIRE, W. J. Cognitive consistency and attitude change. *Journal of Abnormal and Social Psychology,* 1960, **60,** 345–353. (b)

MCGUIRE, W. J. Direct and indirect persuasive effects of dissonance-producing messages. *Journal of Abnormal and Social Psychology,* 1960, **60,** 354–358. (c)

MCNEMAR, Q. *Psychological statistics.* New York: Wiley, 1955.

MORRISSETTE, J. O. An experimental study of the theory of structural balance. *Human Relations,* 1958, **11,** 239–254.

NEWCOMB, T. M. An approach to the study of communicative acts. *Psychological Review,* 1953, **60,** 393–404.

NEWCOMB, T. M. *The acquaintance process.* New York: Holt, Rinehart, & Winston, 1961.

NEWCOMB, T. M. Stabilities underlying changes in interpersonal attraction. *Journal of Abnormal and Social Psychology,* 1963, **66,** 376–386.

OSGOOD, C. E. The nature and measurement of meaning. *Psychological Bulletin,* 1952, **49,** 197–237.

OSGOOD, C. E. Cognitive dynamics in the conduct of human affairs. *Public Opinion Quarterly,* 1960, **24,** 341–365.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning.* Urbana: Univer. Illinois Press, 1957.

OSGOOD, C. E., & TANNENBAUM, P. H. The principle of congruity in the prediction of attitude change. *Psychological Review,* 1955, **62,** 42–55.

PILISUK, M., Cognitive balance and self-relevent attitudes. *Journal of Abnormal and Social Psychology,* 1962, **65,** 95–103.

ROSENBERG, M. J. An analysis of affective-cognitive consistency. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm, *Attitude organization and change.* New Haven: Yale Univer. Press, 1960.

ROSENBERG, M. J., & ABELSON, R. J. An analysis of cognitive balancing. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm, *Attitude organization and change.* New Haven: Yale Univer. Press, 1960.

ROSENBERG, M. J. HOVLAND, C. I., MCGUIRE, W. J., ABELSON, R. P., & BREHM, J. W. *Attitude organization and change.* New Haven: Yale Univer. Press, 1960.

RUNKEL, P. J. Equilibrium and "pleasantness" in interpersonal situations. *Human Relations,* 1956, **9,** 375–382.

SAMPSON, E. E., & INSKO, C. A. Cognitive consistency and performance in the autokinetic situation. *Journal of Abnormal and Social Psychology,* 1964, **68,** 184–192.

SECORD, P. F., BACKMAN, C. W., & EACHUS, H. T. Effects of imbalance in the self-concept on the perception of persons. *Journal of Abnormal and Social Psychology,* 1964, **68,** 442–446.

STEVENS, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1950. Pp. 1–49.

STOUFFER, S. A. An overview of the contributions to scaling and scale theory. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen, *Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950.

SUCHMAN, E. A. The intensity component in attitude and opinion research. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen, *Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950.

SUPPES, P. On an example of unpredictability in human behavior. *Philosophy of Science,* 1964, **31,** 143–148.

TAGIURI, R. Social preference and its perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior.* Stanford, Calif.: Stanford Univer. Press, 1958. Pp. 316–336.

TAGIURI, R., BLAKE, R. R., & BRUNER, J. S. Some determinants of positive and negative feelings in others. *Journal of Abnormal & Social Psychology,* 1953, **48,** 585–592.

WALKER, HELEN M., & LEV, J. *Statistical inference.* New York: Holt, 1953.

WYLIE, RUTH. *The self concept.* Lincoln: Univer. Nebraska Press, 1961.

ZAJONC, R. B. Balance, congruity and dissonance. *Public Opinion Quarterly,* 1960, **24,** 280–296.

(Received March 22, 1965)